DOCUMENT RESUME

ED 059 804 RC 005 960

AUTHOR Spolsky, Bernard; And Others
TITLE A Computer Assisted Study of the Vocabulary of Six-Year-Old Navajo Children. Navajo Reading Study Progress Report No. 9.
INSTITUTION New Mexico Univ., Albuquerque.
SPONS AGENCY Bureau of Indian Affairs (Dept. of Interior), Albuquerque, N. Mex.; Ford Foundation, New York, N.Y.
PUB DATE Aug 71
NOTE 66p.

EDRS PRICE MF-$0.65 HC-$3.29
DESCRIPTORS *American Indians; *Child Language; *Computational Linguistics; Graphemes; *Language Research; *Navaho; Program Descriptions; Sentence Structure; Word Frequency; Word Lists

ABSTRACT

In keeping with the objective of the Navajo Reading Study, to investigate the feasibility and effect of teaching Navajo children to read their own language first, it was decided that more needs to be known about Navajo children and the language they know. Thus, between October 1969 and June 1970, 22 adult Navajo interviewers recorded free conversations with over 200 6-year-old Navajo children at 10 locations on the Navajo Reservation. Interviews were transcribed, in normalized orthography, by one transcriber and key-punched for computer processing. From the total of 11,128 sentences processed, the complete sample of 52,008 words (tokens) represented 8,775 different words (types). Output of the processing included (1) a number of statistical measures, (2) complete concordance giving sentence context, (3) a list of all the words in alphabetical order giving frequency and range, (4) a list of all the words in alphabetical order from the end of the word, (5) a frequency listing, and (6) a number of lists according to various spelling patterns. A concordance giving English loan words in the sample in the context of the sentence in which they occurred was also produced. It should be noted that word lists resulting from this will be used as a filter in preparing reading material for 6-year-old Navajo children. The body of the report provides a description of the study; the appendix includes translated extracts of the interviews and samples of program output. Related documents are ED 035 484, ED 043 004, ED 043 005, ED 043 413, and ED 048 584. (NQ)

ED 059804

RECEIVED
DEC 27 1971
NMSU
E.R.I.C.

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
OFFICE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL OFFICE OF EDUCATION POSITION OR POLICY.

A COMPUTER ASSISTED STUDY OF THE VOCABULARY OF SIX-YEAR-OLD NAVAJO CHILDREN

Bernard Spolsky
Wayne Holm
Jonathan Embry

Navajo Reading Study Progress Report No. 9
The University of New Mexico, August 1971

ED 059660

RECEIVED
DEC 27 1971
NMSU
E.R.I.C.

NAVAJO READING STUDY

The University of New Mexico

August 1971

The work reported herein was supported in part by the United States Department of the Interior (Bureau of Indian Affairs, Navajo Area Office, Contract No. NOO C 1420 3462), and in part by the Ford Foundation.

The symbol chosen for the Navajo Reading Study is an adaptation of one of the graphically elegant Rainbow Protector figures used in the Coyote Chant.[1] Coyote is always a dynamic figure, clever and powerful. When he gets too greedy or vain or arrogant he is killed. But he always returns to life to triumph again. The chant itself, as told by Yoh Hatrali in 1931, tells how the people got corn from the Coyote People, along with the "all important ritual which will keep man safe from harm and in proper balance with the great forces of nature."[2] In the myth the grand-nephew of the Leader of the Yellow Corn People is a hunter. Over near Debeh-entsah he got lost and found himself in the land of the Coyote People. Here he was given four ears of corn, white, yellow, blue and black and as he was given each one he was told, "When you get back to earth take that ear of corn with you and plant it and you will always be able to live on it and increas[3]

---

(1) Slide F 23 #4 Pana-vue slide 2-6, "The Coyote Chant", from the Museum of Navajo Ceremonial Art, Santa Fe.

(2) "The Myth and Prayers of the Great Star Chant and The Myth of the Coyote Chant", recorded by Mary C. Wheelwright, Navajo Religion Series, Vol. IV, page 105. Museum of Navajo Ceremonial Art, Santa Fe. 1956.

(3) Ibid. p. 99.

Background

Started in 1969, the central purpose of the Navajo Reading Study is to investigate the feasibility and effect of teaching Navajo children to read in their own language first. It will thus take its place with a small, but growing body of research into the hypothesis that children can learn to read a second language more easily and with better results if their first introduction to the reading process has been through the medium of their native language.

Before this hypothesis can be tested, it is of course necessary to have available good methods and materials for teaching reading in Navajo. Our first step, therefore, was to make a survey of all Navajo reading materials we could find and to prepare an analytical bibliography (Navajo Reading Study Progress Reports 3 and 7). Material available for teaching reading in Navajo falls into several groups. A number of small books have been prepared by the Wycliffe translators working under the auspices of the Summer Institute of Linguistics. Still in print are a 68-page beginning reader and a set of charts and teaching aids, and a few readers with portions of the bible. A second group of materials are those prepared for the BIA in 1940-50: four bilingual "Little Herder" readers, a bilingual "Little Man's Family" series, and a bilingual "Navajo Life Series". While

these are usable, the religious bias of the former groups and the emphasis on English of the latter limit their suitability considerably. A third group of materials that is just starting to appear is being prepared in connection with the pilot bilingual education programs at Rough Rock Demonstration School, Rock Point Boarding School, and the Bicultural Kindergarten Project. So far, only a few small readers have been published.

To learn more about the possibility of these materials and the methods used to teach reading in Navajo, we organized a conference of Navajo educators with experience in the area (Navajo Reading Study Progress Report No. 6). Each teacher had developed his or her own approach, using whatever materials could be found or developed. Each was thus a pioneer in establishing methods and materials, but no one would suggest he had the method or enough material.

The survey and meeting made clear that before we could make a fair test of the value of teaching reading in Navajo, we would need to see that good materials were available. As a first step to this end, we realized that we must find out more about Navajo children and the language they know.

For a number of years, there has been controversy about the best way of teaching reading in English, focused

especially on what Jeanne Chall has called the "great debate". In simplest terms, the question is whether to emphasize the presentation of the written words as whole units or whether to emphasize the alphabetic nature of the writing system. However one feels about the question, its importance has been to focus attention on the nature of the reading process. A view that we accept is that reading is the process used in translating from alphabetic symbols (meaningless in themselves) to the form of language which is meaningful to the native speaker.

The implications of this are important. Learning to read is learning that writing is a specially coded form of speech. It would seem highly probable that the most effective materials for a child to learn to read with will be closely related to the language he already speaks. This is, of course, the justification for the central hypothesis that the Study is intended to investigate: it was thus completely fitting that before we set out to prepare materials for learning reading that we study the language of the children themselves.

Reflecting their frequent claim that the spoken language is primary and the written language a reflection of it, linguists who became involved with reading not unnaturally argued that beginning reading instruction should be based

on the fact that a child has already learned to talk. From this, they generally went on to argue for strict linguistic controls in the development of the beginning reading materials. Without committing ourselves to methodological questions, however, we still accepted the value of some detailed knowledge of the language of Navajo children.

Approach and limitations.

In recent years, there have appeared a large number of studies of the language of children. A number of possible models were thus suggested. We could choose to study the language of one child in depth, or to survey the language of a number of children. We could choose to base our study on a corpus of free speech, to elicit specific linguistic forms, or to test for comprehension of specific forms. We could choose to collect and analyze our material, or establish a program of ongoing collection and analysis. We needed to decide whether to concentrate on phonology, syntax, lexicon, semantics, or style, or to attempt to cover all of these.

A meeting with a group of Navajo language experts and linguists gave us a chance to explore the values of these various approaches (see Navajo Reading Study Progress Report No. 4). Considering their suggestions and the possibilities

of the Study, we finally decided to base our first investigations on a corpus of six-year-old speech. This corpus would permit a study of grapheme frequency and grapheme patterns, of lexical frequency, of morphological complexity, and a start on some syntactic analysis.

The limitations of the corpus must be emphasized. First, much depends on the interview itself. Our sample was collected in free conversations between adult Navajos and one or more six-year-old children; in only a few cases did we manage to record free conversation between children. It is necessary to recognize that the child's speech in such a situation will not display the full power of his command of language, and may reflect inhibition in the interviewing situation. From the texts, however, we believe that our interviewers were generally successful in overcoming this inhibition.

A second limitation of a corpus is that it does not expose potential. The presence of a word or structure is evidence that a child knows it, but its absence is not evidence that it is not known. Therefore, it must be noted that the various descriptions and lists are not in any way exhaustive. The lists can be used to justify including a word on them in a reader as a word children know,

but the absence of a word or form from the list is not evidence it is not known. A later study will collect data on word availability and possibly word familiarity.

The third limitation of the corpus is that it involves only the topics covered in the interviews. There are a number of domains that have probably not been fully sampled, and one suspects that the interviews reflect the interviewer's beliefs about the children's interest. The planned word availability study will fill in the gaps.

The present study then has been largely limited to the collection of a corpus of the speech of a number of six-year-old Navajo children, and to an analysis of the lexicon. As will come clear, a good deal of phonological and grammatical information has been gained in the process, but this has not been our main present concern.

Our primary aim has been to establish a beachhead: to get a first picture of the language of these children. But there have also been some very practical aims. The word lists produced by the study will be used as a filter when we start preparing reading material for six-year-old Navajo children: if a word occurs in the sample with reasonable frequency and range, a curriculum writer need

have little hesitation in using it. Second, the word frequency list will give indications of words that might well be included in early reading materials. Finally, the whole corpus and the lists derived from it have provided data for a study of the present orthography and its usefulness for writing Navajo.

What we did then was this. Between October 1969 and June 1970, twenty-two different adult Navajo interviewers recorded conversations with over two hundred six-year-old Navajo children at ten different locations on the Navajo Reservation. All the interviews were transcribed, in normalized orthography, by one transcriber and key-punched for computer processing. Altogether, a total of 11,128 sentences were processed. The complete sample consists of a total of 52,008 words (tokens), representing a total of 8,775 different words (types). Output of the processing includes a number of statistical measures, a complete concordance giving sentence context, a list of all the words in alphabetical order giving frequency and range, a list of all the words in alphabetical order from the end of the word, a frequency listing, and a number of lists according to various spelling patterns. Also produced was a concordance giving English loan words in the sample in the context of the sentence in which they occurred.

Collection of the Sample

The interviewers were all Navajo adults with experience working with Navajo children. One of them is an interpreter for a school board; the others are generally teachers, teacher aides, or students in training for teaching. Their instructions were minimal: they were told of the purpose of the study, and of our desire to gain samples of children's speech. No attempt was made to establish interview protocols, to delimit topics, or to prescribe techniques.

Had we at the time had knowledge of the study by Wepman and Hass (1969) of the oral language of English five-, six-, and seven-year-olds, we might have tried to emulate their technique. In their work, they interviewed a smaller number of children (thirty for each age group). Each child was given the sex-appropriate twenty-card array of the Thematic Appreciation Test, and asked to tell a story for each card. There was no prompting by the examiner after the practice session. The influence of the interviewer's language was thus minimized.

It might be argued however that this technique would give a much more limited and perhaps more artificial style than that we obtained in our free-ranging conversations.

Some of our texts were in fact collected by giving the children pictures and asking them to tell the story in their own words. The pictures used were from a set of color photographs of a day in the life of a young Navajo child (Crowder 1970).

The interviewers generally selected the children to be recorded, depending on situation, setting, and availability. The only restriction was that the children should be six years old and able to speak Navajo, although a few older and younger children are in fact included in the sample. The children come from a number of different parts of the Reservation, so that the sample covers the possibility of dialectal variation; some of this may have been lost in the transcription process. A good number of the interviews were at Rock Point Boarding School: there were others at Crystal Boarding School, Cottonwood Day School, Lake Valley Boarding School, Valle Vista School, and Dzilth-Na-O-Dith-Lhe Boarding School. Some of the interviews were made in classrooms: a good number were made in offices and teacher's lounges. Other tapes were made in homes in these areas: Newcomb, Bloomfield, Monument Valley and Shiprock.

A wide range of topics was covered. Some interviewers had the children talk about books or objects in the classroom. A good number had the child talk about home and family life. There was a set of questions on color in one group, and a half dozen children who talked about coyote tales. A couple of children talked about science concepts. A dozen told of field trips. Six told stories to go with the Crowder pictures. Four talked about a picture of a space ship. And several were permitted to talk about what they liked. In some cases, pairs or small groups of children were permitted to converse freely without the intervention of the interviewer.

Some idea of the interviews is given by the translated extracts which appear as Appendix A.

Transcription

An early decision we made was to transcribe in normalized spelling rather than to attempt a phonemic transcription. Our policy was to follow the Young and Morgan orthography with the modifications accepted by the 1969 Conference on Navajo Orthography, and following the computer conventions established by Professor Oswald Werner.

The Conference on Navajo Orthography was called in 1969 by the Center for Applied Linguistics on the

recommendation of an earlier Planning Conference for a Bilingual Kindergarten Program for Navajo Children, which had suggested that a single orthography be adopted for use in the BIA educational system. The general principles on which the 1969 Conference's decision was based were as follows:

1. Consistency is required for pedagogical reasons, but flexibility must be allowed for some time in such matters as shapes of letters, spelling patterns and punctuation. It is premature to expect standardization.
2. The system should be based on the needs of the Navajo speaker, not the non-Navajo.
3. Some consideration should be given to effects of transfer to and from English orthography, but it is more important to base the orthography on the facts of Navajo.
4. Generally, the orthography should attempt to maintain a close phoneme-grapheme fit, but study should be made of tendencies to morphophonemic spelling.

After a review of the various systems that have been used, it was agreed by the Conference to use the modified "Government System", as established by Robert Young and William Morgan. The alphabet agreed on was as follows:

a

b

ch

ch'

d

dl

dz

e

g

gh

h

hw

i

j

k

kw

k'

l

ɬ

m

n

o

s

sh

t

t'

tɬ

tɬ'

ts

ts'

w

x (if 'h' ambiguous)

y

z

zh

'

The prosodic markers to be used were:

1. Length indicated by doubling letter.
2. High tone indicated by acute accent above letter; low tone unmarked.
3. Nazalization indicated by hook under letter.

Among spelling conventions accepted was a decision not to write initial glottal stops.

In order to use this orthography in a computer, certain modifications are necessary: these had been worked out by Professor Oswald Werner for his computerized studies.

1. Upper case letters must be used throughout.
2. The following letters are changed:

   ɬ is written LH

   tɬ is written TL

   tɬ' is written TL'

3. High tone is indicated by a "7" after the vowel.
4. Nasality is indicated by an "8" before the vowel.
5. Syllabic 'n's are written with a following "I".

With our decision to use this orthography and to normalize spelling of the texts, a great deal of importance was necessarily given to the selection of a transcriber. We were fortunate in being able to obtain the services of Babette Holliday, who had been trained in transcribing Navajo by Kenneth Begishe for Professor Werner's project.

To start, Miss Holliday typed the text from the recorded tapes. However, rather than go through a separate stage of key punching, it was later decided to have her do this. Once she had learned to operate the IBM key punch machine, she was able to transcribe directly from tape onto IBM punch cards.

The text was printed from cards and returned to Miss Holliday for editing and correction. As will be readily understood by anybody who has been involved with transcription of this nature, a number of slips and systematic errors have continued to turn up, but basically, the transcription has turned out to be of very high quality.

Editing

As the text, word lists, and concordance have been worked with over the past year, a number of errors in transcription have been noted, and corrections made in the final versions.

The Problem of Word Division

A central problem in our work was computer recognition of words. The problem is that for our purposes, we were forced to accept a definition of a word as something written with blanks on either side of it. This leads to treating BIGHAN 'his house', BIGHANIDI 'at his house', and SHIGHAN 'my house' as separate words.

An effect of this, in our study of word frequencies, is to have a large number of low frequency words. There are two main factors contributing to this in Navajo. First is of course the complexity of the Navajo verb. We made no attempt to deal with this: Werner (1966)

will help make clear our decision. The second is concerned with the use of affixes.

Consider for example the following words which appear in our texts:

	Occurences
O7LTA'	33
O7LTA'DAH	1
O7LTA'DI	12
O7LTA'DIDAH	1
O7LTA'DISH	1
O7LTA'DO7O7	1
O7LTA'D8E7E7	1
O7LTA'GI	3
O7LTA'GO7NE'	2
O7LTA'GO7NE'E7	3
O7LTA'GO7O7	10
O7LTA'GO7O7SH	2
O7LTA'JI7	2
O7LTA'J8I'	12

What we clearly have here is a single word O7LTA, which occurs 84 times, 33 without and 41 with various suffixes. If we could find a way of handling the suffixes, we could get a better picture of the frequency.

In the texts, there are several kinds of affixes:

1. Enclitics:

   KIN 'his house'

   KIN-DI 'at his house'

2. Prefixed possessive pronouns:

   NAALTSOOS 'book'

   BI-NAALTSOOS 'his book'

3. Prefixed post positions:

   DEEYA7 'he will go'

   BEE-DEEYA7 'he will go by means of it'

A further problem is raised by the fact that our transcriber often wrote words together that are separated in Young and Morgan, writing DOO A7T'E7EDAH where they would use DOO A7T'E7E DA.

Enclitics

Some enclitics are suffixed; others are written separately. (See Hale (1965) for a list of enclitics, Young and Morgan (16-20) for information on word status.) All postpositional enclitics are suffixed. Most 'sentence' enclitics are written as words. By convention, DOODA is a single word but negation as three DOO ____ DA.

Sandhi occurs with many, but not all enclitics.

DOO A7T'E7E DA doo+át'é+da

CH'I7NI7L8I7I7GO7O7 Ch'ínlį́į́+góó

NAAT'A7A7NII NE7EZDI Naat'áánii Néez+di

With some words an extra vowel occurs between the word and the enclitic.

HOOGHANIDI hooghan+di

The same enclitic may occur in a reduced form.

HA7A7D8E7E7'SH hááděę́'+shą'

Or it may assimilate to the preceding vowel:

SIDA7A7SH sidá+ísh

Two or more enclitics may occur together.

O7LTA'DII7SH ółta'+di+ísh

These may collapse and assimilate.

KINTAHGO7O7SH kintah+góó+ísh

(Werner gives an example, borrowed from Hale, of a sequence of three enclitics.)

Possessive Pronouns

Possessive pronouns are prefixed to the noun. There are some instances of assimilation: BA'A7LHCHI7NI7 bi-'ářchíní "his/her children" affecting the form of the pronoun, or the noun: BE'EWE7E' bi-'awéé' "his/her baby". Secondary possession is marked somewhat differently. BE'ABE' is "her (indefinite) milk" bi-'a-be'; BIBE' would be "her own milk (from her own breast)".

The initial letters of possessive pronouns resemble those of a number of postpositions. The latter might be analyzed as a sequence of pronoun + postpositional element. Thus

BI- 'his' BIKA7A7' on him

NI- 'your' NIKA7A7' on you

SHI- 'mine' SHIKA7A7' on me

The prefixed possessive pronouns resemble both the independent possessive and independent subjective pronouns.

BI- 'his' BI7 'he, his'

NI- 'your' NI 'you, yours'

SHI- 'my' SHI7 "I, mine'

They also resemble the object pronouns of some verb forms.

BI- 'his' BIDINI7 'you say it to him'

NI- 'your' NILHNI7 'he says it to you'

SHI- 'my' SHIDINI7 'you ask me'

The NI- resembles the object pronoun NIHI- 'us two', 'you two':

NI-'your' NIHILHNI7 'he says it to you two'

It also resembles a number of potentially verb-initial verb-prefixes, as NI7LHMA7A7S 'you roll it', NI7LHMA7A7Z 'I rolled it', ni-imperfectives, ni-perfective, and a number of other forms. Some nouns require a high tone possessive. Thus, NI7LA' 'your hand', not *NILA'.

The SHI- resembles the SHI- of the similar prefix of si-perfective verbs with "esh"-initial and or final stem: SHIJAA' 'they are', SHIBÉÉZH 'boiled', SHI7NI7SHISH 'you poked it'. Finally, the possessive pronouns resemble the initial syllable of a number of words which are in no way related: e.g. BILH 'sleep', NI' 'earth', SHIID 'ssst!'.

Postpositions

Postpositions are usually written as separate words. A few are prefixed to the verb.

Postpositions may be thought of as consisting of two parts: a pronominal and a postpositional element. The postpositional element may be written immediately after a noun, the noun in place of the pronominal element. TSIN BITAH SIZ8I7 'woods-them-among-he stands', TSINTAH 'woods-among'.

Verb phrase postpositions are usually written as separate words. YAA NAAGHA7.

Often a form will be normalized as two words even though phonologically there is but one: /yeiíní'á/ YAA YI7NI7'8A7.

Some forms are written as a single word. BIK'IIDEESHWOLH might be read as bik'i 'ideeshwol if written BIK'I IDEESHWOLH (bik'i yideeshwol).

And some forms are written as a single word because they are, in a sense, idiomatic; the meaning of the sum is more

than that of the parts. BEE DEEYA7 'he's coming by means of it', but with BEE 'ELD8OOH the derivation is obviously bee, 'by means of it, adildǫǫh, shooting occurs'. However, BEE'ELD8OOH is not felt to be 'by means of it shooting occurs', but 'gun'!

Postpositions can occur with enclitics. BITAHDI 'at a place among them'.

Thus, postpositions may occur as words, as prefixes or as suffixes.

Processing of word division.

Against this background, we now consider the process of editing the text to achieve appropriate word division. There are three possible approaches: visual inspection, computer-controlled partial matching, and computer-controlled affix stripping.

a) By visual inspection is meant the examination and marking of forms one by one. Only visual inspection will ensure near-complete accuracy. But even this is difficult: relatively few Navajos can do it and then only for a few hours at a time. The texts contain approximately 9,000 tokens. To treat each word as a unique problem, when many are instances of the same problem, seems extremely wasteful of time and effort.

b) By partial matching is meant using the computer to locate pairs of words that match over a predetermined portion or area of their length.

Thus partial matching might help locate words which are essentially the same except for different prefixes or suffixes.

HOOGHAN-D8E7E7' 'house-from'

HOOGHAN-J8I' 'house-up to'

BI-ZHE7'E7 'his-father'

SHI-ZHE7'E7 'my-father'

There are a number of problems with this approach. It may divide words that we do not want to treat as separate words. Thus, it may separate verb prefixes from stems:

NA7SH-'8I7 'I usually fix it'

NA7SH-CHAH "I usually cry'

NA7SH- is the verb prefix marking first person singular in the iterative mode--not something we would wish to treat as a separate word. And, it leaves as words the verb stems -'8I7 and -CHAH neither of which we wish to treat as words. Or we may find stems divided from prefixes: DAA-NI7 'they say', NA7A7NE7I7LH-NI7 'he says it again'. -NI7 is a verb stem having to do with 'saying'. Not only does partial matching treat the stem as a word but the prefixes are also treated as words. This fault might be remedied by insisting that we must be able to match both portions of the word we wish to divide.

Partial matching, with the requirement that both portions must be matched, still allows one to obtain instances where the one portion is a common recurring verb prefix and the other a common recurring stem. Thus, DEESH-CHA might occur because DEESH- (marking first person singular, future) might be found in: DEESH'A7A7LH, DEESHBISH, DEESHCH'ALH, DEESHDAH, and -CHA 'to cry' in YISHCHA, YI7CHA, NA7SHCHA, GHO7SHCHA.

This might be remedied by insisting that <u>both</u> portions also occur as free forms. This requirement, however, would be too strong for it would exclude all but those compounds whose components occurred freely. It would handle only those words which we used to separate, and these only if both words were already in the corpus. Thus, we might wish to separate a form YAANAAGHA7 to read YAA NAAGHA 'he's doing'. If our requirement is that both words already occur as words in the text, and YAA did not happen, then our computer would not separate. It is even less likely that affixes would occur as free forms. Thus, we might wish to mark a form BINAALTSOOS to read BI-NAALTSOOS 'his book'. But if the computer cannot find a word BI in the text, it will be unable to divide the word. Some words, such as body parts, occur only as possessed forms. For the possessed form SHIJAAD 'my leg' there is no free form *JAAD. Thus, the requirement that both elements be found would cause the computer to fail to divide correctly the word SHI-JAAD.

Partial matching then fails as a technique for word-division.

By affix stripping is meant using the computer to set off given affixes from the words to which they are attached. One may wish to begin with a list of affixes and automatically set them off. A number of problems result. A major one is that there are any number of forms homographic with some of the affixes. Thus, the computer may recognize the possessive pronoun BI- in the word BIIZHII and mark it *BI-IZHII.

There are some affixes homographic with verb stems. Thus, our transcriber consistently writes -DAH as a negating element. -DAH represents several homophonous verb stems, and several other words. Affix separation would incorrectly separate these. Thus, while the computer would correctly separate the transcribed form *DOO A7TE7EDAH to read DOO A7TE7E DA 'it isn't', it would incorrectly separate a form like DEESHDAH 'I'll wipe it off' to read *DEESH DAH.

Similarly, one affix may be included within a longer one; the possessive pronoun BI- is included in any number of postpositions. Affix stripping would incorrectly separate the shorter form. Thus, the form BIKA7A7'DAH'ASDA7HI7

might be incorrectly separated as *BI-KA7A7'DAH'ASDA7HI7 Failure to separate the word as *BIK'A7A7' DAH 'ASDA7HI7 precludes the possibility of correct separation on a second pass BIK'A7A7 DAH "ASDA7HI 'chair'. These problems might be handled by insisting that the non-affixed element occur as a free word. At first blush, this combination of the affix-stripping and partial matching seems to offer the best hope of properly dividing words. In practice, however, it runs into all kinds of trouble, basically because the non-affixed element does not necessarily occur in exactly the same form, or because it may not occur in the corpus. Thus, the combination of the two requirements is too strong. It fails to mark any number of real divisions. Some examples follow:

Word initial glottal stops. In earlier orthographies, vowels never occurred at the beginning of words or syllables. In more recent orthographies, the word-initial glottal stop is omitted. This may cause problems. The form *SHIKA7'ANILYEED should be divided as SHIKA7 ANILYEED 'help me' but the computer will attempt to separate SHIKA7 leaving 'ANILYEED; it will be unable to find a free form with a glottal stop inital (the free form, if it occurs, will occur without a glottal stop, as ANILYEED). This

might be remedied by allowing a match of a 'V-initial form with a V-initial word.

Extra Vowels. HOOGHANIDAH should be separated (and corrected) as HOOGHANI DA the i appears to be epenthetic). But the program will not find a word HOOGHANI. One cannot very well allow an additional optional consonant unless one can be sure that these enclitics do not occur with words that differ only by a single vowel.

Geminate Vowels. AT'E7EDAH should be separated (and corrected) to AT'E7ED DA 'girl-not' but the program will not be able to find a word AT'E7E7 or AT'E7 lengthened by some other enclitic. One cannot very well allow additional optional consonants between the stem and suffix because verb stems may differ from each other only by the presence or absence of a final consonant.

Tone Sandhi. DOO SIDA7ADAH should be separated as DOO SIDA7A DA, but, unless SIDA7A occurs as a free form, we are unable to do so. We might allow an optional V. This may not cause any problem with V7V syllables as stems do not seem to occur in isolation.

The recurrent dilemma here is whether allowances one makes to obtain the correct division of words won't also allow unforeseen incorrect divisions. Even if solutions are found to these problems, some problems would remain.

Thus, [illegible] the convention that one writes BIKA7A7'ADA7NI7 'table' but BIKA7A7' DAH 'ASDA7HI7 'chair'. The same BIKA7A7 occurs in both forms. This is simply a spelling convention partly analagous to SHIPROCK but CHURCH ROCK as English place names on the Reservation. One can handle this problem only by visual inspection or resort to a dictionary.

Relativals also cause real problems. Should they be counted? With verbs they seem to serve to make nouns. Thus, one could argue that HATAALH and HATAALHI7 are related but different words, as are their equivalents 'sing' and 'singer' in English. But they merely relativize nouns. One would probably want to accept SIO' 'stars' and S8O'I7GI7=7 'the ones that are stars' as instances of the same word.

The distinction between nouns and verbs cannot be made by computer at this time. A dictionary hook up would be required. But even if one were able to hook up the computer to a dictionary, and to make the proper adjustments for glottal stops, geminate consonants, and sandhi, there is no complete solution.

But we did not have a computer hooked up to a dictionary. And we did not have a working set of rules for adjustment to glottal stops, geminate consonants, or sandhi phenomena. Given plenty of time, one might work out a number of these problems by trying out an adjustment rule, visually inspecting the results and modifying the rule, until by progressive refinement one gets more adequate rules. But we did not have this sort of time, and there is some question whether the results would justify the effort. Given the shortness of time we might write loose adjustment rules, incorrectly marking a number of words and thereby obtaining speciously high frequencies for some forms; or we might write tight adjustment rules, letting a number of words go unmarked and thereby obtaining speciously low frequencies. Or, we could abandon the effort to write adjustment rules for the computer and go back to visual inspection; which is what we have done.

Our first task was to prepare a list of potentially separable words or affixes. In doing this, we were able to make use of a reverse alphabetized listing of the words in the text, which of course placed suffixes together. A program then printed out all words containing these elements and alongside them, put one or more tentative

divisions. Navajo research assistants checked these lists visually, deciding which agreed with the conventions we had laid down.

As is so often the case, we found that it took too much time and effort to have a computer do all the work, but finished up with man-machine cooperation that let each fill its best role.

Computer Processing.

Computing was carried out at the University of New Mexico's Computing Center, using an IBM 360/40 until June 1970, and after that an IBM 360/67. The programs were written by Jonathan Embry in PL/I-F. The programs used up to 102 k storage, need two 800 bpi tapes, and one 2314 disk.

From the taped interviews, the text was transcribed onto IBM cards in a free format that allowed as many blanks as desired between sentences or words. Control cards were used to separate and identify interviews. Each speaker was assigned a three character code which was punched at the start of each sentence. These cards were then built onto a disk file by a program that reformatted the material into fixed-length records. Each record contained one sentence (maximum length of 300 characters), along with identifying information. Each sentence was assigned a number

indicating its relative position within the text. This number, along with speaker identification for each sentence is printed out in the TEXT LISTING. The material used consisted of 21 interviews, containing 11,128 sentences spoken by 218 children and 14 adults.

The concordance was generated by building a file containing each word used, along with the appropriate sentence identifying number. When these were sorted alphabetically, the number was used to retrieve the sentence (s) that contained each word. For every word used, each sentence containing that word was printed, along with the speaker code and sentence number (in case it is necessary to refer to TEXT to get the full context of the sentence) for that sentence.

From the same word file, it was possible to produce both a straight alphabetical list (sorted left to right) and a reverse word list (sorted right to left), each containing every word used in the text.

In addition, a file was created that contained each unique word, along with frequency and range figures. From this were printed frequency lists, showing frequencies and ranges for various words and letter patterns.

A program similar to the first concordance program printed entries for selected words only. This was used to produce a concordance of English loan words.

A major problem turned out to be getting similar occurences of the same word to be listed together. After trying several techniques (partial matching, specific rules, etc.), manual correction by native speakers proved to be the most expedient.

List of computer programs used.

1. Card reformatting.
2. Text listing.
3. Separate sentences into words.
4. Concordance.
5. Alphabetical frequency listing.
6. Reverse word listing.
7. Enclitic stripping.
8. Partial matching.
9. Grapheme and unit frequency count.
10. Text correction 1.
11. Text correction 2.
12. Spelling list.

References

1. Crowder, Jack L. Stefanii doo Ma'ii. 1970.

2. Hale, Kenneth. The Distribution of the Class II Prefixes in Navaho. Unpublished M.A. thesis, Indiana University, 1956.

3. Murphy, Penny, editor. Teaching Initial Reading in Navajo: Report of a Conference of Educators held at Kayenta. Navajo Reading Study Progress Report No. 6. 1970.

4. Navajo Reading Study Progress Report No. 4, 1969.

5. Spolsky, Bernard, Agnes Holm, Penny Murphy. Analytical Bibliography of Navajo Reading Materials. Navajo Reading Study Progress Report No. 3, 1969. Revised, Navajo Reading Study Progress Report No. 7, 1970.

6. Young, Robert and William Morgan. The Navajo Language. 1945.

7. Wepman, Joseph M. and Wilbur Hass. A Spoken Word Count. 1969.

APPENDIX A

Translated Samples of the Interviews

The following excerpts are translations of the original Navajo text. Each sentence is numbered, and followed by the speaker's code. Interviewers are identified by numerical codes (001) and children by initials (OJR). Any words that were in English in the interview are given in quotation marks. Notes are added in square brackets.

1000 001 What's your name?

1001 OJR Juanita Ration.

1002 001 How old are you?

1003 OJR "six".

1004 001 Where are you from?

1005 OJR From home.

1006 001 What name do they call your place?

1007 OJR Far away.

1008 001 Who do you usually stay with at home?

1009 OJR My mother and my father.

1010 001 Who else?

1011 OJR Dennison and Billy.

1012 001 Who are they?

1013 OJR Those that are going to school at "West Mes .

1014 001 What does your father usually do?

1015 OJR He doesn't do anything.

1016 001 He just sits around?

1017 OJR Yes.

1018 001 What about your mother?

1019 OJR She's weaving.

1020 001 What does your big brother do at home?

1021 OJR He doesn't do anything.

1022 001 You mean he also doesn't do anything?

1023 001 What about your sister?

1024 OJR She doesn't do anything.

1025 001 What do you usually do?

1026 OJR I don't do anything.

1027 001 You mean everybody just sits around?

1028 001 How big is your house?

1029 OJR Not too big.

1030 001 How many rooms does it have?

1031 OJR "three".

1032 001 What do you have in the kitchen?

1033 OJR dishes.

1034 001 What else?

1035 OJR "cups".

1036 001 What else?

1037 OJR that's all.

1038 001 You only have dishes and cups?

1039 001 Do you have a stove?

1040 OJR We don't have any.

1041 001 What do you cook your food with?

1042 OJR We don't cook our food.

1043 001 What are in the other two rooms?

1044 OJR "bedroom".

1045 001 What else?

1046 OJR just one.

1047 001 What is in that one?

1048 OJR "bed".

1049 001 What else?

1050 OJR blankets.

1051 001 What else?

1052 OJR That's all.

1053 001 What is in the other room?

1054 OJR clothes and "shirt".

1055 001 What else?

1056 OJR "bed".

1057 001 There's another bed in there?

1058 001 What color of paint is your house painted with?

1059 OJR five.

1060 001 What's five?

1061 001 Do you have any sheep?

1062 OJR none.

1063 001 What about horses?

1064 OJR none.

1065 001 What about cattle?

1066 OJR none.

1067 001 What about a car?

1068 OJR none.

1069 001 What do you travel around in?

1070 001 On foot?

1071 001 What does your grandmother usually do?

1072 OJR She doesn't do anything.

1073 001 What about your grandfather?

1074 OJR He's gone.

1075 OJR My father took my grandfather somewhere.

1076 001 When?

1077 OJR One time.

1078 001 What happened?

1079 OJR He didn't come back last night.

1080 001 What about your father?

1081 001 Your father came back?

1082 OJR Yes.

1083 001 What about your mother?

1084 OJR She came back.

1085 001 Where did they go?

1292	002	What is your name?
1293	ORM	"Rebecca Marie".
1294	002	How old are you?
1295	002	What is this town called?
1296	002	Do you understand Navajo?
1297	002	What do you call this place?
1298	002	How do you come here in the morning?
1299	ORM	"Bus".
1300	002	Do you come everyday?
1301	002	What is your teacher's name?
1302	002	Did you go to school last year?
1303	ORM	No.
1304	002	Where did you attend school?
1305	002	What did you learn about?
1306	002	Right here?
1307	ORM	Right at home.
1308	002	Do you have a mother?
1309	002	What is her name?
1310	ORM	"Marie".
1311	002	Do you have a father?
1312	002	What do you usually do here at school?
1313	002	What do you usually do around here?
1314	002	Do you have an older sister?
1315	002	How about your younger sister?

1316	ORM	I have "three".
1317	002	What are their names?
1318	ORM	"Eva" and "Andrea" and "Maxine".
1319	002	Where are they now?
1320	ORM	They are over at home.
1321	002	They are at home.
1322	002	Who are they staying with.
1323	ORM	My mother and "Martha".
1324	002	Who is "Martha".
1325	002	What do you usually do when you travel on the "bus"?
1326	002	Does the "bus" come for you?
1327	002	Do you walk?
1328	ORM	No.
1329	002	How do you arrive at school?
1330	ORM	They drive us here.
1331	002	In what?
1332	ORM	"Bus".
1333	002	What do you usually do when you are on your way to school?
1334	ORM	I usually look at a "book".
1335	002	Where do you eat at noon?
1336	ORM	Over there.
1337	002	Do you eat here?
1338	002	What do you eat?
1339	002	What is your name?
1340	OHF	"Henry Francisco".

1341	002	Where do you live?
1342	OHF	Where the grey cow is standing.
1343	002	You live around here.
1344	002	Do you live in town?
1345	002	Do you have a mother?
1346	OHF	She is called "Lilly".
1347	002	How about your father?
1348	OHF	"Henry"
1349	002	Does your father work?
1350	002	Where?
1351	OHF	At "Gallup".
1352	002	How about your mother?
1353	OHF	My mother also works.
1354	002	Where does she work?
1355	OHF	I don't know the place.
1356	002	Do you have older brothers?
1357	002	What are their names?
1358	OHF	"Calvin" and "Emanual" and "Nelton" and "Arnold".
1359	002	How about sisters?
1360	OHF	You mean the younger ones?
1361	002	Your younger sisters.
1362	OHF	Yes.
1363	002	You have some?
1364	002	What are their names?
1365	OHF	"Karen" and "Jean".

1366	OHF	That's all.
1367	002	Do you have younger brothers?
1368	OHF	No.
1369	OOU	When did you enroll in school?
1370	OOU	When did you come here?
1371	OOU	Did you come to school as a beginner?
1372	OOu	Do you like school?
1373	OOU	What do you do in class?
1374	OOU	Tell about what you do.
1375	OHF	I usually play.
1376	002	What else do you do?
1377	002	Who is your best friend in school?
1378	OHF	I don't have one.
1379	002	How about the girls?
1380	OHF	None of them, either.
1381	002	You play by yourself.
1382	OHF	Just a few.
1383	OHF	Only four.
1384	002	You play with them.
1385	002	Do you have fun playing?
1386	002	What do you usually do at noon?
1387	002	Do you understand the word noon?
1388	002	When you go there at noon, what do you do there?
1389	OHF	You mean the eating place?
1390	OHF	Not me, I don't eat over there.

1391 OHF I usually eat at home.

1392 002 Do you live nearby?

1393 002 Do you usually walk to school every day?

1394 002 When don't you come to school?

1395 OHF I don't know.

1396 002 Do you miss school some days?

1397 OHF No, I attend classes.

1398 002 You go to school every day. Right?

1399 OHF My brother also goes to school.

1400 002 Where does your brother go to school?

1401 OHF Over that direction.

1402 002 You mean he goes to another school.

1403 OHF He goes to school here.

1404 002 Does he help you?

1405 OHF I have another one younger but we are the same height.

1406 002 Does he go to school, too?

1407 002 Where does he go to school?

1408 OHF Over there.

1409 OHF Over that way. On the other side.

1410 002 What is his "teacher's" name?

1411 OHF I don't know.

1412 002 What do you usually do at home early in the morning?

1413 OHF I put on my clothes.

1314 002 What else do you do?

1315 OHF I come here with my brothers.

1416	002	How do you help your mother?
1417	OHF	I don't help her.
1418	OHF	She works.
1419	002	Who takes care of you when your mother works?
1420	OHF	My father and my "Auntie".
1421	002	She lives with you?
1422	OHF	Just her, only "Auntie".
1423	002	Do you help her with the housework?
1424	OHF	We take out the trash.
1425	002	Can you name some fourlegged animals?
1426	002	Which ones do you know?
1427	OHF	"Horse".
1429	002	Horse and what else?
1430	002	How about those animals with white faces and they have horns?
1431	OHF	"Cow".
1432	002	How do you say it in Navajo?
1433	OHF	Cow.
1434	002	How about those animals that stay around homes and sometimes bark? What are they called?
1435	OHF	Dog, Cat.
1436	002	What else?
1347	002	There is one that crawls on the ground like this. What is that one called?
1348	002	It crawls like this. What is it?

1439 OHF "Snake".

1340 002 Can you say it in Navajo?

1441 OHF Snake

1442 002 Tell me some more about four-legged animals.

1443 002 How many do you know?

1444 002 Can you name some pack animals?

1445 002 Those that you can ride.

1446 002 Do you understand what I'm talking about?

1447 OHF "Horse"

1448 002 Say it in Navajo.

1449 OHF Horse.

1450 002 How about those animals that have white faces with horns?

1451 OHF Cow.

1452 002 What about those animals that stay around homes and bark? What are they called ?

1453 OHF Dogs.

1454 002 Again, repeat it.

1455 002 Cats?

1456 002 Can you name some more animals?

1457 002 What do the Navajos herd?

1458 002 They're white.

1459 OHF Sheep.

1460 F01 What are the ones that are similar to sheep but they have horns. What are they called?

1461 F01 They are the same size as the sheep.

1462	FO1	Sometimes these animals are milked. What are they?
1463	FO1	They are in the same herd with sheep.
1464	OHF	Rams.
1465	002	There are others that are similar, with horns or no horns and in the same herd with sheep. What are these animals called?
1466	002	Can you think of those little animals that are usually along side the mature sheep?
1467	002	They're white.
1468	002	Do you know it?
1469	002	Do you sometimes herd sheep?
1470	OHF	Just my maternal grandfather herds sheep.
1471	002	Do you help him sometimes?

2060 004 "Phillip Ned" will tell us about the growth of a bean plant.

2061 OPN We brought in some red dirt first.

2062 OPN And then we put some black dirt in too.

2063 OPN And then we put in a bean.

2064 OPN And then we watered it.

2065 OPN And then we put it by the window.

2066 OPN Several days later, we opened up a bean. It was still little.

2067 OPN Several days after that, we opened up another bean it had grown bigger.

2068 OPN Several days later, we opened up another, and found that the roots had started growing.

2069 OPN After several days, we opened up another, there were lots of roots that had grown.

2070 OPN We opened up another bean, there were even more roots.

2071 OPN And then the leaves came out.

2072 OPN It grew because the sun was shining on it and it was warm in the room and we watered it as needed and we took good care of it.

2073 004 Leon, what part of a lettuce do we eat?

2074 OLN We eat the leaves.

2075 004 How about carrot? "carrot"

2076 OLN We eat th root.

2077 004 How about celery? "celery"

2078 OLN We eat the stem.

2079 004 How about onions?

2080 OLN We eat the root.

2081 004 How about potatoes?

2082 OLN We eat the root.

2083 OLN We put in some red dirt and some black dirt.

2084 OLN Then we dug a hole for it and put it in.

2085 OLN And then we covered it with dirt.

2086 OLN And then we watered it.

2087 OLN Several days later, we opened up another one and it grew bigger.

2088 OLN Several days later, we opened up one and it had roots.

2089 OLN Several days later, we opened up another and there were more roots and they grew longer.

2090 OLN The potato grew bigger.

2091 OLN Several days later, we opened up another one, it grew more roots.

2092 OLN There were lots of roots.

2093 OLN And it had leaves.

2094 OLN We watered it and the sun was shining on it and it was on the windowsill.

Navajo Reading Study Interview No. 10

2736	004	The "Beginners" will tell you about animals.
2737	002	The fox eats food.
2738	004	What is their food?
2739	002	Grass, sheep, meat, and they build their homes.
2740	OZZ	The ducks swim in water.
2741	ZZZ	Some animals live in it.
2742	ZZZ	Rabbits.
2743	OOX	Some animals walk.
2744	OOX	Animals walk.
2745	OXX	Some cats have homes made for them.
2746	XXX	Fish swim in water.
2747	OOY	Some animals aren't good pets.
2748	OOY	Dogs are pets.
2749	OYY	Some animals fly.
2750	OYY	Some birds fly.
2751	YYY	Some animals fly.
2752	YYY	Birds fly.
2753	OOW	Dogs eat meat.
2754	OWW	Some animal babies look like their mothers.
2755	OWW	Some are born not looking like their mothers.
2756	OWW	Frogs lay eggs and then just run away.
2757	OWW	The tadpoles first grow their hind legs by eating weeds under water.
2758	OWW	And then their front legs grow again.

2759	OWW	Then they are shorter.
2760	OWW	Then they turn into toads. [toads or frogs]
2761	WWW	Cats eat rats.
2762	OGB	My name is "George Benally".
2763	OGB	Mice moved into an all ready home that were already there.
2764	OVJ	My name is "Vernor Jim".
2765	OVJ	Homes are made for horses.
2766	OOZ	Some animals have homes made for them.
2767	OOZ	Homes are made for cows.
2768	OLC	My name is "Lorraine Coggeshell".
2769	OLC	Some animals hop.
2770	OLC	Rabbits hop.
2771	ODB	My name is "Dennis Begay".
2772	ODB	Some animals move into natural homes.
2773	ORS	Bears move into homes made by nature.
2774	ORS	My name is "Ramona Sideburn".
2775	OKH	Some animals carry their homes on their backs.
2776	OKH	My name is "Kimberly Holly".
2778	OKH	Rabbits make burrows and live in them.
2779	OFG	My name is "Felisita Gatewood".
2780	OFG	Some animals have their homes made for them.
2781	OFG	Homes are made for sheep.

3704 003 What about school, what do you do there?

3705 003 Are you learning?

3706 OGB Yes.

3707 003 What are you learning?

3708 OGB Some things from books.

3709 003 Really?

3710 OGB Yes.

3711 003 Tell me about that, what kind of things do you learn?

3712 003 Do you write, too?

3713 OGB No.

3714 003 Why?

3715 OGB Our teacher just gives those "pencils" to us.

3716 003 Your teacher gives them to you?

3717 OGB Yes.

3718 003 And what else?

3719 OGB Then we keep them.

3720 003 What did you do during Christmas?

3721 003 What did tney send you during Christmas?

3722 OGB Papers, all this size. [The object is a magic slate with cellophane or plastic.]

3723 OGB Some of them were very long.

3724 OGB This narrow,

3725 OGB the papers were this narrow.

3726 OGB Those shiny things on them, there were two, one on top of another.

3727 OGB And the paper is this long.

3728 OGB It is this wide.

3729 OGB It is spread inside and we write on it; some red things, they're rounded like this.

3730 OGB It has a hole in it like this.

3731 OGB We dip it in water, and we draw anything right here, that's what they sent us.

3732 003 Really?

3733 OGB Yes.

3734 OGB When you lift it up, it erases itself.

3735 003 Oh, yes, those.

3736 003 And what else?

3737 OGB That's all.

APPENDIX B

Samples of Program Output

SENTENCE SPEAKER

8527 ORG AAD0707 DI7I7 ASHKII SIT8I7H.

3379 ORW AAD0707 ERASER T'A7A7 BI7LA7TAHDI DAHSI'8A7AGO NI7LE7I7GO7NE' NA'ADZI7I7LHHAAL.

8299 OTT AADO707 KODI LHA' DAHJIZDA7.

2274 OVT BIDA7A7JI7 EELWOD EII AAD0707 NI7LE7I7GO7NE' BAA EELWODA7 NI7LH'8I7.

9855 OXX AAD0707 NI7LH'8I7 MONKEY.

9831 OXX FIL8A7A. T'A7A7 AAD0707 YILH YADOOLTALH.

9877 OZZ AAD0707 NI7LH'8I7 K80707 SI'8A7.

9858 OZZ AAD0707 FIRE YICH'8I' CALL I7I7LE7E7H YA'S

9887 OZZ NI7LE7I7 TS17DLI TAAHJIJEEH NI?T'8E7E7' TA7YI' YI'OLH AAD0707.

4179 002 AAD0707 T'0707 BA7HA7DZIDI7GII A7T'E7EGO DAHISK'ID8E7E YIKA7A7'GO7DEI HAASTA7A7L.

4129 002 AAD0707 HA7LA7TSOHI7GI7I7 T'A7A7 A7LHAH KO7T'E7EGO KODI HA7LA7YA7ZHI7 DAHYOOLE7LHI7GI7I7 BI7IGHAHDI KO7T'E7EGO NE7IIDIILE7E7H.

4522 002 AAD0707 KO7T'E7EGO JIDILHTS'8O'.

4504 002 AAD0707 KO7T'E7EGO T'A7A7LHA7H.7DI BINI7NA7A7JI7LHT'IH.

4196 002 AAD0707 KWE'E7 NI7T'8E7E7' GAH ALHTSOH I7I7Y8A7A7' NI7T'8E7E7' T'AH NI7T'8E7E7' K80707 CH'E7E7H DIGHA7HII BAANA7A7DIILWOD.

1808 003 AAD0707 DIDZE7 NE7E7LH A7JI7LE7EGO BAANI7ZDIIT'A7A7ZH.

1528 003 AAD0707 LH8I7I7' LHIBA7HI7GI7I7DO7' NA7A7NA7 LHA' JINI7.

3271 003 T'0707 DAHSOHKE7EGO AAD0707 T'A7ADOOLE'E7 NI7LE7I7 O7LTA'GO7NE' HA'A7T'I7SH8I7I7 BAANAAH'AASH. NI7LE7I7 HOOGHANDIDO7'.

6090 003 T'0707 AADD707 SOOKE7.

7126 003 AAD0707 HODITINA'GO A7A7D8E7E7' NIHAANI7DIISHDA7A7LH.

9615 010 T'A7A7 AAD0707 NEHI7SO7OT8A7H.

THIS WORD WAS USED 11 TIMES BY 3 INTERVIEWER(S).
THIS WORD WAS USED 71 TIMES BY 30 CHILDREN.
THIS WORD WAS USED 82 TIMES BY 33 SPEAKERS ALTOGETHER.

ACTIVITY

7621 OVB DO707 SWING BEE A?SHT'8I7I7 NI7T'8E7E7', DO707 NI7LE7I7 GO7NE' YAH'ANA7SHO' ACTIVITY ROOM GO7NE' DO707 CH'I7NI7NA7SHWO'.

THIS WORD WAS USED 0 TIMES BY 0 INTERVIEWER(S).
THIS WORD WAS USED 1 TIMES BY 1 CHILDREN.
THIS WORD WAS USED 1 TIMES BY 1 SPEAKERS ALTOGETHER.

SAMPLE - Alphabetical Word List

STEM WORD	INTERVIEWER FREQ.	INTERVIEWER RANGE	CHILDREN FREQ.	CHILDREN RANGE	TOTAL FREQ.	TOTAL RANGE
	29	3	21	6	50	9
ATSA7	0	0	1	1	1	1
BA7	1	1	0	0	1	1
HWE7E7LDI	1	1	0	0	1	1
NIMA*A7LHCHI7NI7	1	1	0	0	1	1
OF	0	0	2	2	2	2
SHILHNI7	0	0	1	1	1	1
*YA7A7H	0	0	1	1	1	1
----------DD7D7	0	0	1	1	1	1
*	1	1	1	1	2	2
*AP	0	0	1	1	1	1
*A7HWIINI7T*8I7*I7	0	0	1	1	1	1
*A7HWIINI7T*8I7*I7GI7I7	0	0	1	1	1	1
*A7LNE7E7H	0	0	1	1	1	1
*A7T*I7I7DD7*	1	1	0	0	1	1
*A7T*I7SH8I7I7	0	0	1	1	1	1
*A7TIIN	1	1	0	0	1	1
*DI*NI7	1	1	1	1	2	2
BI-*DI*NI7	1	1	0	0	1	1
BI-*DI*NI7-I-GO	0	0	1	1	1	1
*DI*NI7I7	0	0	1	1	1	1
*DCONIID	0	0	1	1	1	1
*E7*8E7E	1	1	1	1	2	2
E7E7	8	5	29	10	37	21
BI-*E7E7*	8	5	21	10	29	15
SHI-*E7E7*	0	0	8	7	8	7
*E7E7TSOH	0	0	1	1	1	1
*E7E7TSOH8E7E	0	0	2	2	2	2
*IIH	0	0	1	1	1	1
*IISHT*O*	0	0	1	1	1	1
*I7SH	1	1	0	0	1	1
JIILHDL8A7A7	0	0	1	1	1	1
JOOSOLA7A7	0	0	1	1	1	1
*NAA	0	0	4	2	4	2
*NII*NII*	0	0	1	1	1	1
*NIIGH8A7A7	0	0	1	1	1	1
*NIIGH8A7A7*I7SH	0	0	1	1	1	1
*NIILHHX8I7	0	0	1	1	1	1
*NIILHH8I7	1	1	4	3	5	4
*NIILHH8I7I7SH	0	0	1	1	1	1
*NIITS8A7	0	0	1	1	1	1
O7LTA	1	1	0	0	1	1
*O7LTA*I7	1	1	0	0	1	1
*8AA*I7YIILAA	0	0	1	1	1	1
A	2	2	8	7	10	9
A*	1	1	1	1	2	2
A*AAH	0	0	1	1	1	1
A*A7A7N	0	0	1	1	1	1
A*A7A7NGQ7YAA	0	0	1	1	1	1
A*A7T*I7I7	1	1	0	0	1	1
A*DEESHDLALH	0	0	1	1	1	1
A*E7E7*	0	0	1	1	1	1
A*HOOT*E7	1	1	0	0	1	1
A*NOOLHKALH	0	0	2	2	2	2
AA	1	1	3	3	4	4

SAMPLE - Frequency Word List

CHILDREN FREQUENCY LISTING PAGE 1

WORD	INTERVIEWED FREQ.	INTERVIEWED RANGE	CHILDREN FREQ.	CHILDREN RANGE	TOTAL FREQ.	TOTAL RANGE
E7I7	702	13	1127	60	1829	73
A7A7-00707	192	12	1016	60	1208	72
00707	189	14	722	110	911	124
LA7	145	10	481	60	626	70
T'A7A7	311	12	441	60	752	72
A0C'	24	8	424	72	448	80
NI7T'8F7E7'	137	10	376	70	513	80
DI7I7	356	11	370	60	726	71
NI7LE7I7	144	9	350	60	494	69
LE7I'	17	5	339	60	356	65
LHEH	283	13	326	64	609	77
T'07D7	74	10	318	76	392	86
K80707	60	7	312	54	372	61
SHI7	35	6	287	60	322	66
DA	188	10	285	79	473	89
LHA'	75	11	257	60	332	71
KWE'E7	131	7	249	58	380	65
SH81717	257	9	246	60	503	69
BILH	81	10	235	60	316	70
Y8F7F	35	4	217	44	252	48
JINI7	178	3	200	29	378	32
DOO	131	10	195	60	326	70
BEE	248	11	189	64	437	75
J07	75	7	157	44	232	51
I7NDA	65	7	149	32	214	39
LH81717'	102	8	147	42	249	50
YA'	114	5	145	28	259	33
A7T'E7	131	9	144	51	275	60
KODI	11	6	129	42	140	48
YA77HI7	4	3	127	36	131	39
E7IYA7	91	10	126	42	217	52
T'A7ADBOLE'F7	61	7	112	43	173	50
BA7	72	8	105	34	177	42
A7KO	123	9	100	27	223	36
LHF7F7CH8AA'I7	7	5	99	27	106	32
GA'	1	1	92	38	93	39
ALHDO7'	66	8	86	41	152	49
NIDA	1	1	83	36	84	37
HO7LAH	1	1	82	33	83	34
BAA	131	12	81	40	212	52
NI7LH'8I7	10	2	81	26	91	28
NA7A7NA7	38	8	80	32	118	40
TO7	3	2	79	45	82	47
GO7NF'	51	7	77	47	128	54
EI	38	3	75	37	113	40
YILH	11	7	75	37	86	44
SHI-MA7	7	3	74	34	81	37
DIBE7	9	6	73	40	82	46
ASHKII	13	5	71	20	84	25
A7ADI	56	6	68	29	124	35
DI7IGII	20	2	68	16	88	18
HA'A7T'I7I7	472	13	68	36	540	49
SH8A'	582	14	68	35	650	49
YILWOLH	9	3	68	31	77	34
ALHTSOH	19	4	67	35	86	39
BI-MA7	14	5	66	30	80	35

COMPUTING CENTER

A7T*E7HI7GI	A7T*E7HI7		
A7T*E7HI7GII	A7T*E7HI7		
A7T*E7II	A7T*E7II		
A7T*E7IILA7	A7T*E7II	A7T*E7IILA7	LA7
A7T*8I7	A7T*8I7		
A7T*8I7GII	A7T*8I7		
A7T*8I7I7	A7T*8I7I7		
A7T*8I7I7LA7	A7T*8I7I7	A7T*8I7I7LA7	LA7
BAA*I7I7JIL	BAA*I7I7JIL		
BAA*I7I7JILGO	BAA*I7I7JIL		
BAADAA*IHGO	BAADAA*IH		
BAADAA*IHLE7I*	BAADAA*IH	BAADAA*IHLE7I*	LE7I*
BAADIILWOD	BAADIILWOD		
BAADIILWODLA7	BAADIILWOD	BAADIILWODLA7	LA7
BAANA*ALDEEH	BAANA*ALDEEH		
BAANA*ALDEEHLA7	BAANA*ALDEEH	BAANA*ALDEEHLA7	LA7
BAANANINA7A	BAANANINA7A		
BAANANINA7AGO	BAANANINA7A		
BAANITSA7HA7KEES	BAANITSA7HA7KEES		
BAANITSA7HA7KEESGO	BAANITSA7HA7KEES		
BAANI7SI7ST*8IIDLA7	BAANI7SI7ST*8IID		
BAANI7SI7ST*8IIDNI*	BAANI7SI7ST*8IID	BAANI7SI7ST*8IIDNI*	NI*
BABY	BABY		
BAEYDO7*	BABY		
BA7HA7DZID	BA7HA7DZID		
BA7HA7DZIDGO	BA7HA7DZID		
BA7HA7DZIDI7GII	BA7HA7DZIDI7		
BE*ELYAA	BE*ELYAA		
BE*ELYAALA7	BE*ELYAA	BE*ELYAALA7	LA7
BE*ELYAAY8E7E	BE*ELYAA		
BEAR	BEAR		
BEARDO7*	BEAR		
BEELDLE7I7	BEELDLE7I7		
BEELDLE7I7Y8E7E	BEELDLE7I7		
BEES	BEES		
BEESDO7*	BEES		

	INTERVIEWER FREQUENCY	INTERVIEWER RANGE	CHILDREN FREQUENCY	CHILDREN RANGE	TOTAL FREQUENCY	TOTAL RANGE
AL8A7AJ8I*	8	4	4	3	12	7
BIL8A7AJ8I*	1	1	1	1	2	2
HAANI7ZA7ADJ8I*	1	1	0	0	1	1
ANA7*A7T*A7H8E7EJ8I*	0	0	1	1	1	1
KWI7Y8E7EJ8I*	1	1	0	0	1	1
LHAHJ8I*	1	1	1	1	2	2
A7LHAHJ8I*	2	2	3	3	5	5
NAHJ8I*	11	2	1	1	12	3
BINAHJ8I*	0	0	3	1	3	1
BI7LA7TAHJ8I*	0	0	1	1	1	1
A7NI7ZAHJ8I*	0	0	1	1	1	1
BIGH8A7AHJ8I*	1	1	0	0	1	1
BIK*IHJ8I*	0	0	7	6	7	6
A7K*IHJ8I*	0	0	1	1	1	1
DOOT*A7ALHJ8I*	1	1	0	0	1	1
ADOOT*A7ALHJ8I*	1	1	0	0	1	1
DOOLEELHJ8I*	1	1	0	0	1	1
ALHTSOHJ8I*	1	1	0	0	1	1
BIDI7I7LHD8ASHJ8I*	0	0	1	1	1	1
HOOGHANJ8I*	0	0	1	1	1	1
DAH*ALZHINJ8I*	0	0	1	1	1	1
KOJ8I*	35	5	22	9	57	14
AYA7HA7J8I*	2	1	0	0	2	1
HO7NI7ZA7HA7J8I*	0	0	1	1	1	1
DIBE7J8I*	1	1	0	0	1	1
HO7NI7ZA7HI7J8I*	0	0	1	1	1	1
NAAZNILI7J8I*	1	1	0	0	1	1
HA7NI7J8I*	1	1	0	0	1	1
KE7HAT*I7NI7J8I*	0	0	1	1	1	1

SAMPLE - Enclitic Stripping

O7OLYE7EDAH	O7OLYE7E DA
NA'NITINDAH	SHINA'NITIN DA
ISII'DAH	SHITSII' DA
7DAH	SHI7 DA
I7IGODAH	SH8I7IGO DA
SIDA7ADAH	SIDA7A DA
SIK'AZDAH	SIK'AZ DA
SINDAH	SIN DA
SITS'ILDAH	SITS'IL DA
SIT8A7ADAH	SIT8A7A DA
SWINGDAH	SWING DA
T'A7ADOOLE'E7DAH	T'A7A7DOOLE7'E7 DA T'A7A7 OOOLE7'E7 DA
T'O7ODAH	T'O7O DA
TA7A'DAH	TA7A' DA
TA7A7'TS'A7ADAH	TA7A7'TS'A7A DA TA7A7'TS'A7A DAH
TA7DADOOGIZHDAH	TA7DADOOGIZH DA
TE7LIIDAH	TE7LII DA
TI7DZI7'AHI7DAH	TI7DZI7'AHI7 DA
TL'I7ZI7DAH	TL'I7ZI7 DA
TL'I7ZI7YA7ZHI7DAH	TL'I7ZI7YA7ZHI7 DA
TL'O7O'DIDAH	TL'O7O'DI DA
TO7NANEESDIZI7GO7ODAH	TO7NANEESDIZI7GO7O DA
TO7SHCHI7I7NDAH	TO7SHCHI7I7N DA
TS'IILZE7I7DAH	TS'IILZE7I7 DA
TS'I7DA7DAH	TS'I7DA7 DA
TSEE8I7I7HTS'A7ADAH	TSEE8I7I7HTS'A7A DA TSEE8I7I7TS'A7A DAH
TSE7HOOTSOOI7GO7ODAH	TSE7HOOTSOOI7GO7O DA
TSIHALHDAH	TSIHALH DA
TSINDAH	TSIN DA
TSI7DIIDAH	TSI7DII DA
TSI7LHKE7IDAH	TSI7LHKE7I DA
YA'DAH	YA' DA
YAANAA'AASHI7DAH	YAANAA'AASHI7 DA
YAANAAGHA7ADAH	YAANAAGHA7A DA
YAANAAGHA7HA7DAH	YAANAAGHA7HA7 DA
YAANAAKAII7DAH	YAANAAKAII7 DA
YAANITSE7KEESDAH	YAANITSE7KEES DA
YAH'I'I7I7T'I'DAH	YAH'I'I7I7T'I' DA
YA7'A7SH8O7ODAH	YA7'A7SH8O7O DA
YA7DAALHTI'DAH	YA7DAALHTI' DA
YA7JI7LHTI'DAH	YA7JI7LHTI' DA
YA7JI7LTI'DAH	YA7JI7LTI' DA
YA7LHTI'DAH	YA7LHTI' DA
YA7NI7LHTI'DAH	YA7NI7LHTI' DA
YA7SHTI'DAH	YA7SHTI' DA
YA7ZHI7DAH	YA7ZHI7 DA
YE7E7HO7SINDAH	YE7E7HO7SIN DA
YIDIDOOTS'8I7ILHDAH	YIDIDOOTS'8I7ILH DA
YIDIITS'A'DAH	YIDIITS'A' DA
YIDZAASDAH	YIDZAAS DA
YIDZAAZDAH	YIDZAAZ DA
YIIHIYIINI7ILHDAH	YIIHIYIINI7ILH DA
YIINI7ZINDAH	YIINI7ZIN DA
YIK'I'DILYE7EDAH	YIK'I'DILYE7E DA
YINI7DAALDZIDAH	YINI7DAALDZI DA
YIT'8I7IDAH	YIT'8I7I DA
YI7I7NI7'8I7IDAH	YI7I7NI7'8I7I DA
YI7I7NI7TA7ALDAH	YI7I7NI7TA7AL DA
YI7I7NI7ZINDAH	YI7I7NI7ZIN DA
YI7LHTA'DAH	YI7LHTA' DA
YO7'EELWODI7DAH	YO7'EELWODI7 DA
YSE7EDAH	YSE7E DA
ZHE7'E7DAH	ZHE7'E7 DA

SAMPLE - Grapheme and unit frequency count

RELATIVE FREQUENCIES

DI-GRAPH	ADULTS	CHILDREN	TOTAL	DIFFERENCE
CHX	0.0000	0.0000	0.0000	0.0000
DZ	0.0000	0.0000	0.0000	0.0000
GH	0.0000	0.0000	0.0000	0.0000
HX	0.0000	0.0000	0.0000	0.0000
KW	0.0000	0.0000	0.0000	0.0000
SX	0.0000	0.0000	0.0000	0.0000
TL	0.0000	0.0000	0.0000	0.0000
TL'	0.0000	0.0000	0.0000	0.0000
TSX	0.0000	0.0000	0.0000	0.0000
TS	0.0000	0.0204	0.0167	-0.0204
H	0.0000	0.0204	0.0167	-0.0204
ZH	0.0000	0.0204	0.0167	-0.0204
D	0.0000	0.0612	0.0500	-0.0612
T	0.2727	0.0408	0.0833	0.2319
K	0.0909	0.0816	0.0833	0.0093
C	0.0909	0.1020	0.1000	-0.0111
L	0.0000	0.2449	0.2000	-0.2449
S	0.1818	0.2245	0.2167	-0.0427
C	0.3636	0.1837	0.2167	0.1800

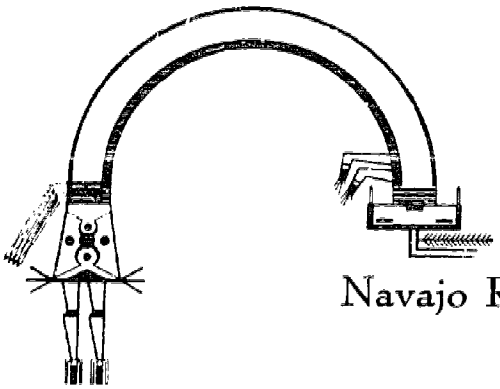
ABSOLUTE FREQUENCIES

DI-GRAPH	ADULTS	CHILDREN	TOTAL	DIFFERENCE
CH	4	9	13	-5
CH'	1	5	6	-4
CHX	0	0	0	0
DL	0	3	3	-3
DZ	0	0	0	0
GH	0	0	0	0
HW	0	1	1	-1
HX	0	0	0	0
K'	1	4	5	-3
KW	0	0	0	0
LH	0	12	12	-12
SH	2	11	13	-9
SX	0	0	0	0
T'	3	2	5	1
TL	0	0	0	0
TL'	0	0	0	0
TS	0	1	1	-1
TSX	0	0	0	0
ZH	0	1	1	-1

SAMPLE - Spelling List

WORDS FOUND CONTAINING 4 SYLLABLES

BITA'I7GI7I7
BITA7I'AHDI
BITA7T'AHD8E7E7'
BITL'AAJI'E7E7'
BITC7SHCHI7NE7E
BITSE7SK'EHGO7O7
BIYA7ZHI7CO
BIYA7ZHI7Y8E7E
BI7DAHAASTL'8O7O
BI7DAHALTAALH
BI7DAHOOLH'AA
BI7DAHWIIL'AA
BI7DANIILDIN
BI7DANI7ZI
BI7DI7LNIHI7
BI7DO7LHNIHI7
BI7HONEEL'8A7A
BI7KAZHNITA7
BI7KA7ZHDI7LWO'
BI7LA7TANJRI'
BI7NA7A7JII'NIL
BI7NA7A7ZHDI7LNIH
BI7NI7OEIILJAH
BI7NI7JI7GA7ASH
BO7HOJIILH'AAH
BO7HONEEDL8I7I7
BO7HWIINILH'8A7A7'
CAYADITTC
CH'AANISOOKAI
CH'AATA7DOOKAI
CH'AHASFKE7EGO
CH'E7'E7LWOD8E7E
CH'E7'E7TIINE7E
CH'E7'E7TIINGI
CH'E7E7DICHA7HII
CH'E7E7SHINI7SID
CH'ILDIILYE7SII
CH'ILDIILYE7SII7SH
CH'I7DAFIISHAN
CH'I7DAYIISHAN
CH'I7HONI7SHO7O7'
CH'I7JI'NI7NIL
CH'I7NA7'A7T'A'
CH'I7NA7A7HA7SHKAAH
CH'I7NA7A7NA7'A7
CH'I7NA7A7NIIKAI
CH'I7NA7DZIDII'
CH'I7NA7LWO'GO
CH'I7NE7IJAHGO
CH'I7NI7DEIIKA7A7H
CH'I7NI7DEIILB8AS
CH'I7NI7NA7SHWO'
CHAHA'OFGI
CHAHALHFFELFGI
CHAHASH'OHGI
CHAHASH'OFGO
CHIDI7YA7ZHI7
CHCOANOOLH'8I7I
CHCOAYOCLH'8I7



Navajo Reading Study • The University of New Mexico

1805 Roma N.E.
Albuquerque, New Mexico 87106
Telephone 505: 277-5749

BIBLIOGRAPHY

Reading Books

1. *Mósíłgai* Story by Marlene Atcitty, Illustrations by Caryl McHarney. 8 pp.

2. *Pábii dóó Mási.* Story by Judy Harvey, Illustrations by Caryl McHarney. 16pp. w/cover. (in preparation)

3. *Jasper Naané.* Story by Irene Silentman, Illustrations by Caryl McHarney. 6 pp. (in preparation)

4. *Táá' Hastóí.* Story by Judy Harvey, Illustrations by Caryl McHarney. 8 pp. w/cover. (in preparation)

Progress Reports

1. *NRS Progress Report No. 1.* R. Paul Murphy. July, 1969.

2. Proposal for Language Analysis Study. *NRS Progress Report No. 2.* August-September, 1969.

3. Analytical Bibliography of Navajo Reading Materials. Bernard Spolsky, Agnes Holm, Penny Murphy. *NRS Progress Report No. 3.* October-November, 1969. [ED 035 484].*

4. Report of a Conference on the Language Sample, held in Albuquerque on December 4-5, 1969. Edited by R. Paul Murphy. *NRS Progress Report No. 4.* December, 1969. [ED].*

5. Navajo Language Maintenance: Six-Year-Olds in 1969. Bernard Spolsky. *NRS Progress Report No. 5.* March, 1970. [ED 043 004].*

6. Teaching Initial Reading in Navajo: Report of a Conference of Educators held at Kayenta, January 30-31, 1970. Edited by Penny Murphy. *NRS Progress Report No.6.* April, 1970. [ED 043 005].*

7. Analytical Bibliography of Navajo Reading Materials. (Revised and Enlarged Edition). Bernard Spolsky, Agnes Holm, Penny Murphy. *NRS Progress Report No. 7.* June, 1970. [ED 043 413].*

8. Literacy in the Vernacular: The Case of Navajo. Bernard Spolsky and Wayne Holm. *NRS Progress Report No. 8.* March, 1971.

9. A Computer Assisted Study of the Vocabulary of Six-Year-Old Navajo Children. Bernard Spolsky, Wayne Holm and Jonathan Embry. *NRS Progress Report No. 9.* August, 1971.

10. A Spoken Word Count of Six-Year-Old Navajo Children. Bernard Spolsky, Wayne Holm and Babette Holliday, Judy Harvey, Marlene Atcitty, Irene Silentman, Jonathan Embry. *NRS Progress Report No. 10.* August, 1971.

11. Navajo Spelling Lists. Wayne Holm. *NRS Progress Report No. 11.* August, 1971.

12. Grapheme and Unit Frequencies in Navajo. Wayne Holm. *NRS Progress Report No. 12.* August, 1971.

13. Navajo Language Maintenance II: Six-Year-Olds in 1970. Bernard Spolsky. *NRS Progress Report No. 13.* August, 1971.

14. Navajo Language Maintenance III: Accessibility of School and Town as a Factor in Language Shift. Bernard Spolsky. *NRS Progress Report No. 14.* August, 1971.

15. Preparing Reading Materials in Navajo. Marlene Atcitty, Judy Harvey Richard Johnsen, Caryl McHarney, Linda Rogers, Irene Silentman. *NRS Progress Report No. 15.* August, 1971.

16. English Loan Words in the Speech of Six-Year-Old Navajo Children. Agnes Holm, Wayne Holm, Bernard Spolsky. *NRS Progress Report No. 16.* August, 1971.

Published Elsewhere

1. "Navajo Reading Study". *English for American Indians.* Spring, 1970, pages 13-19.

2. Analytical Bibliography of Navajo Reading Materials (Enlarged and Revised Edition). BErnard Spolsky, Agnes Holm, Penny Murphy. Curriculum Bulletin No. 10, United States Bureau of Indian Affairs, 1970.

3. Navajo Language Maintenance: Six-year-olds in 1969. Bernard Spolsky. *Language Sciences.* December, 1970, No. 13, pages 19-24.

4. Literacy in the Vernacular: The Navajo Reading Study. Bernard Spolsky. Paper presented to the Council on Anthropology and Education Symposium on Cognitive and Linguistic Studies, 69th Annual Meeting of the American Anthropological Association, San Diego, November 19, 1970. [ED].*

5. Literacy in the Vernacular: The Case of Navajo. Bernard Spolsky and Wayne Holm. *Studies in Language and Linguistics, 1970-71*, edited by Jacob Ornstein. University of Texas at El Paso Press.

6. Bilingualism in the Six-Year-Old Navajo Child. Bernard Spolsky and Wayne Holm. Paper to be presented at the Conference on Child Language, Chicago, November 22-24, 1971.

*Available from ERIC Document Reproduction Service.